

README

Summary

My homework is Resnet50-based, which is the well-known Net in image-classification area. But for better operating in our set I used more tricks (Adaptive learning rate, small batchsize, modify ResNet50, and DataAugumentation)

The simplest way is to execute '**ForTA.py**'. It loads the weights what I trained using the various technique.

How to use

For scoring

Excecute **ForTA.py**. There is no need to use any arguments. It reads "FtanukiCharNet.pth" weight file on the same folder.

Requisites

- Pytorch
- torchvision
- tanukiCharNet.py, tanukiDataAug.py in same folder.

Design

Abstract

For overcoming the small dataset, I adopted 4 techniques.

1. Augumentation.

My codes generates randomly rotated, shifted, and resized figures without loss. In short, different to the common transform functions, the whole character is not destroyed. It can be achieved by boundary extraction. My code extracts the ROI(Region of Interest) using bounding box function. To make the function well-operated, I applied bilateralFilter and adaptive thresholding in prior. And then for denosing, I applied randomly constrast increasing. From ROI, I resized ROI randomly, but using LANCZOS for lower reszing noise. And then translate it in only 350*350 array.

Because of the speed problem, I implemented this by two methods. First one is simple, I made custom transform function for Pytorch. It looks nice and easy to use. But data loading is too slow because numpy cannot support multiprocessing. For solving it, I just implemented the program to save augumented image in HDD first,"gen_aug.py". And then train it with "tanukiChar_trainer_with_gen_aug.py". It is more faster.

2. Small batch size

By some papers, they argued that generalization performance increase in more various learning rate under small batch size. It recommends batch size as 8.

3. Adaptive learning rate

At higher learning rate, we can get more faster learning. But there can be divergence or vibration. So I decrease the learning rate epoch by epoch.

4. Modify ResNet50

Sadly, ResNet50 only supports 224*224 images with 1000 classes. So I adopted AdaptiveAvgPool2d layer and modified the final FC layer to 5 without downscaling of input images.